

雑誌九十種表記表の統計

著者	宮島 達夫
雑誌名	日本語科学
巻	1
ページ	92-104
発行年	1997-05
URL	http://doi.org/10.15084/00001969

雑誌九十種表記表の統計

宮島 達夫
(京都橘女子大学)

キーワード

国立国語研究所, 現代雑誌九十種, 語彙調査, 表記のゆれ

要旨

国立国語研究所は1956年の雑誌九十種について大規模な語彙調査を実施した。その結果は報告書として発表されているが、今回、その各異表記の度数をふくめた、全部の語彙の度数を言語処理データ集7として公表することになった。本報告は、そのデータの統計的な分析である。修正したデータによって、これまでの報告書にのっている統計表を書きなおしたほか、異表記の度数についての統計をあらたに追加し、表記のゆれが多いのは和語・動詞、すくないのは漢語・固有名詞といった点をあきらかにした。

今回、国立国語研究所言語処理データ集7として、『(フロッピー版) 雑誌九十種表記表』を刊行することになった。この報告は、その解説と、そこにおさめられたデータによる雑誌九十種語彙調査の基本的な統計である。

1. 表記表の解説

この調査対象は1956年のものであって、いまとなつては、〈現代〉とよぶには、ふさわしくない。それでも言語処理データ集7を刊行することには、つぎのような意味がある。

- 1) 報告21『現代雑誌九十種の用語用字 第一分冊 総記および語彙表』にのっている語彙表は、度数7までであり、度数1までの全語彙の度数をしるした語彙表は、印刷物の形では発表されていない。マイクロフィッシュには全語彙の度数表がはいっているが、これは利用しやすいとはいえない。
- 2) 言語処理データ集7では、単語の使用度数だけでなく、その表記ごとの度数をもしるした。このような形での表記表は、発表されていない。(なお、ここには「むずかしい〜むつかしい」のような語形のゆれもふくめたが、活用形の差は無視した。)
- 3) すべての見出し語に語種と〈品詞〉とをしるしてあり、4万語の見出し語をもった一種のコンピューター用辞書として利用することが可能である。(ここでいう〈品詞〉については、あとの説明を参照。)

報告21には、助詞・助動詞の度数表もあるが、これは言語処理データ集7では省略した。また、

報告21の語彙表には○●などの符号の度数もふくまれているが、これも今回ははぶいた。また、雑誌は5つの<層>に分類されていたが、言語処理データ集7では、層別の内わけを無視して、各層の合計度数だけをしめた。

以下、調査単位との関係で、注意すべき点をしるす。調査単位のくわしい規定は、報告21、報告25を参照していただきたい。

- 1) 常識的な意味での単語が表記表になく、もっとこまかい単位にわかれていることがある。たとえば、「自動車」「飛行機」「小学校」などの度数があがっていないが、これらの用例がなかったのではなく、それぞれ「自動／車」「飛行／機」「小学／校」と分割して、単独の「自動」などに合併して集計したのである。
- 2) 「入出国」の類は、「入国」1、「出国」1としてある。
- 3) 接辞も単語なみにあつかわれている。それで、「おとうさん」という単語の度数はない。これは「お／とう／さん」という3単位に分割されている。
- 4) 「あそびくらす」「あばれまわる」などは複合動詞として1単位になっている。しかし、生産力のたかい接尾的動詞・形容詞は接辞あつかいであり、「あそび／はじめる」「あばれ／だす」や「あつかい／にくい」などは、それぞれ、きれている。同様に、「読み／方」も2単位であり、このさい、「読み」は動詞としてある。助数詞も「五／人」「八／杯」などと数とは別の単位になる。なお、これら後要素が連濁をおこしていることがあるが、それらは連濁のない形と合併した。たとえば、助数詞「杯」には「ばい」「パイ」などもあるが、これらは「はい」「ハイ」とあわせて数えた。
- 5) いわゆる助動詞のうち、「(ら)れる」「(さ)せる」「たい」は接尾辞と同様に、動詞・形容詞としてあつかっている。
- 6) 可能動詞（「読める」など）は、もとの動詞（「読む」など）に合併してある。
- 7) 接尾辞「さ」のついた形は、派生名詞としないで形容詞に合併してある。

以上のような規定は、この表記表をつかうばあいには、しておく必要がある。たとえば、動詞「読む」の度数には、ふつうの意味での動詞のほかに「読み方」や「読みはじめる」の度数も合算されている。動詞「ぬう」の表記には、「ぬう」や「縫う」のほかに「縫」が15回でてくるが、おくりがなのないこの形は、すべて「縫方」の前半をきった形である。このような異質のものを合併して、「ぬう」のおくりがなを論ずることは危険である。（ただし、動詞でおくりがなのない形には、ほかに「不可侮」のような、漢文の白文にでてくるものもある。）

2. 基本統計

全体の語数は、つぎのとおりである。

ことなり 39,997 のべ 438,760

これらの語は一般語と固有名（人名・地名）にわかれる。はじめに、固有名をのぞいた一般語だけの統計をあげる。

一般語は語種・品詞の観点から分類される。ただし、ここで品詞というのは『分類語彙表』の

[体 (名詞) ・用 (動詞) ・相 (形容詞・副詞) ・その他] の4類である。文章のなかでは、おなじ単語が名詞としても副詞としてもつかわれる、ということがあるが、便宜上1語には1つの品詞名しかあたえなかった。しかも、じっさいの用例をみて品詞を決定したわけではないので、実例は名詞的な用法だけなのに、3という分類になっている、というような可能性もある。

語種が問題になるものの例をあげると、「尼・あばた・瓦・さじ・みそ」などは和語、「絵 (え)」も和語だが、「図絵」は漢語。「カルマ・インモ・ゆず」は漢語とした。

		ことなり	%	のべ	%
(語種)	和語	11,152	36.7	221,399	53.7
	漢語	14,369	47.3	170,159	41.3
	外来語	2,988	9.8	12,462	3.0
	混種語	1,868	6.1	8,286	2.0
(品詞)	体 (名詞)	23,862	78.6	255,151	61.9
	用 (動詞)	3,457	11.4	97,576	23.7
	相 (形容詞・副詞)	2,839	9.3	51,965	12.6
	その他	219	0.7	7,614	1.8
	一般語計	30,377		412,306	

つぎに、固有名統計。「人」は、略称の「Q」のように、人名であることはたしかだが、姓か名か不明のもの、「姓名」は「エノケン」のように、姓と名とを結合したものである。

人名	7,539	18,223
人	4	8
姓	3,371	9,852
名	4,153	8,351
姓名	11	12
地名	2,081	8,231
固有名計	9,620	26,454

さて、うえの語種・品詞の統計は固有名をふくんでいない。固有名は、語種にわければ和語(「広島」)・漢語(「東京」)・外来語(「ロンドン」)・混種語(「お茶の水」)にわかれ、品詞としては名詞である。それで、これを語種としては第5の独自の種類とし、品詞としては名詞に合併して、全語彙を分類しなおすと、つぎのようになる。

		ことなり	%	のべ	%
(語種)	和語	11,152	27.9	221,399	50.5
	漢語	14,369	35.9	170,159	38.8
	外来語	2,988	7.5	12,462	2.8
	混種語	1,868	4.7	8,286	1.9
	固有名	9,620	24.1	26,454	6.0
(品詞)	体(名詞)	33,482	83.7	281,605	64.2
	用(動詞)	3,457	8.6	97,576	22.2
	相(形容詞・副詞)	2,839	7.1	51,965	11.8
	その他	219	0.5	7,614	1.7
全語彙計		39,997		438,760	

つぎに、語種と品詞を関連させて集計する。以下の数字は、固有名をふくまない、一般語だけの範囲のものである。

[全語種]	(ことなり)				(のべ)			
	体	用	相	他	体	用	相	他
和語	6,136	3,263	1,567	186	83,069	95,541	35,489	7,300
漢語	13,355	-	1,002	12	154,544	-	15,350	265
外来語	2,821	4	145	18	12,117	5	302	38
混種語	1,550	190	125	3	5,421	2,030	824	11

[混種語]	(ことなり)					(のべ)				
	体	用	相	他	計	体	用	相	他	計
和+漢	454	1	35	-	490	2,319	2	115	-	2,436
和+外	131	-	-	-	131	241	-	-	-	241
漢+和	480	188	86	3	757	1,857	2,027	704	11	4,599
漢+外	100	-	1	-	101	157	-	1	-	158
外+和	182	1	3	-	186	487	1	4	-	492
外+漢	203	-	-	-	203	360	-	-	-	360

[外来語]	(ことなり)						(のべ)					
	体	用	相	他	計	%	体	用	相	他	計	%
英語	2,294	3	115	8	2,420	81.0	10,376	3	266	23	10,668	85.6
フランス語	155	1	14	1	171	5.7	624	2	19	1	646	5.2
ドイツ語	92	-	4	-	96	3.2	250	-	4	-	254	2.0
イタリア語	40	-	4	1	45	1.5	82	-	5	3	90	0.7
オランダ語	39	-	-	-	39	1.3	284	-	-	-	284	2.3
スペイン語	20	-	1	-	21	0.7	41	-	1	-	42	0.3
ポルトガル語	21	-	1	-	22	0.7	150	-	1	-	151	1.2
ロシア語	23	-	2	-	25	0.8	51	-	2	-	53	0.4
ラテン語	12	-	1	1	14	0.5	25	-	1	1	27	0.2
中国語	19	-	1	3	23	0.8	52	-	1	3	56	0.4
原語不明	54	-	1	3	58	1.9	77	-	1	6	84	0.7
その他	52	-	1	1	54	1.8	105	-	1	1	107	0.9

以下に、いくつかの問題点をしるす。

まず、品詞のうち、用の類（動詞）があるのは原則として和語か混種語であって、「愛する」は混種語、「愛好する」は「愛好」と「する」にわけて2語としたから、漢語や外来語、およびこれらを後要素とする混種語には動詞がないはずである。うへの表で混種語の〈和＋漢〉に動詞が1語あるのは、「白茶ける」をそうみとめたものである。また、外来語の動詞4語というのは、「アー(are)、アーム(I'm)、カム(come)、セ(c'est)」である。その他の類は接続詞・感動詞・陳述副詞などであって、漢語には「即、万歳、勿論」などがある。外来語については疑問があるかもしれないが、「アーメン、イエス、バイバイ、ハオハオ」などの感動詞である。

つぎに外来語の原語について。〈原語不明〉としたのは、形からいって外来語とおもわれるが、原語の見当のつかない、以下のようなもので、このなかには日本製の、つまり外来語の定義にあわないものがあるかもしれない。

アドルム(薬品)、アバババ(酒場)、アルコ(洗剤)、カコ(ストロボ)、
 ジーニゼイ(回教神話にでてくる悪霊)、黄色大駱駝(チョンサリック)、
 モアナ(楽団名)、ラモラ(架空の会社)

外来語の原語で〈その他〉としたのは、以下の54語である。

アイヌ語(アイヌ、イヨマンテ、オットセイ、昆布、鮭、シシャモ、シャモ[和人]、トド、ラッコ)
 アラビア語(デストゥール[憲法])
 インドネシア語(マンゴ)
 エスペラント(エスペラント、ナップ、プーポ、ロマーナ)
 オランダ語＋英語(コーヒーカップ、コーヒーフェアー、ゴムカタン、ゴムテープ、ゴムバンド、
 ゴムボート、ゴムロール、マドロスパイプ)

カンボジア語 (アンコールワット)
 ギリシャ語 (アルファー, μ)
 サモエード語 (プヤシンカ [黒い木=石炭])
 タイ語 (シャモ [軍鶏])
 タガログ語 (カソイ [くるみの木の一種], サンパギタ [植物名], タガログ)
 チベット語 (達頼 <ダライ>, ラマ)
 ドイツ語+フランス語 (アルサロ)
 ドイツ語+英語 (エチレンオキサイド, ディーゼル・カー, バイト・ガイド)
 フランス語+英語 (コッペパン, パレエシューズ, ピコミシン, ブークレーヤーン, ロココ・スタイル)
 ヘブライ語 (メシア, アーメン)
 ポルトガル語+英語 (ボタンホール)
 英語+オランダ語 (パイプオルガン, フレオンガス, フロントガラス, ベビーオルガン)
 英語+フランス語 (オペラバレエ)
 英語+ポルトガル語 (カフスボタン)
 朝鮮語 (オンドル, チョンガー, ホアン [園])

以上の数値は、報告書にでているものと、すこしちがうところがある。純粹のかぞえちがいもあつたし、語種などの認定を訂正した箇所もある。しかし、そのちがいは、全体からみれば、ごくちいさい。語種の分類について、報告書の数値と比較してみよう。／のまえが報告書にある数、あとが今回の数で、() のなかは%である。

	ことなり		のべ
和語	11,134 (36.7) / 11,152 (36.7)	221,875 (53.9) / 221,399 (53.7)	
漢語	14,407 (47.5) / 14,369 (47.3)	170,033 (41.3) / 170,159 (41.3)	
外来語	2,964 (9.8) / 2,988 (9.8)	12,034 (2.9) / 12,462 (3.0)	
混種語	1,826 (6.0) / 1,868 (6.1)	8,030 (1.9) / 8,286 (2.0)	

したがって、たとえば、よく引用される報告25のp.61にあるグラフを、数値以外の部分について修正する必要はない。

3. 度数別の統計

報告25にならって、度数別の段階をつけ、統計をとると、つぎのようになる。報告書には固有名詞をはぶいた数値がのっているが、今回はこれをふくめたものを「総計」とした。「計」は一般語を合計したものであり、語種・品詞それぞれの合計とおなじである。

ことなり

	和	漢	外	混	名	動	形	他	計
1	4482	5409	1567	1059	10287	1179	995	56	12517
2	1748	2325	485	298	3880	476	464	36	4856
3- 4	1574	2194	393	236	3395	519	456	27	4397
5- 8	1167	1656	255	128	2451	424	306	25	3206
9-16	780	1161	173	76	1646	311	215	18	2190
17-32	551	768	73	37	1045	220	146	18	1429
33-64	376	461	27	19	593	161	116	13	883
65-	474	395	15	15	565	167	141	26	899
計	11152	14369	2988	1868	23862	3457	2839	219	30377

	人	地	総計
1	4280	1269	18066
2	1420	305	6581
3- 4	1034	232	5663
5- 8	528	132	3866
9-16	200	72	2462
17-32	61	46	1536
33-64	15	16	914
65-	1	9	909
計	7539	2081	39997

のべ

	和	漢	外	混	名	動	形	他	計
1	4482	5409	1567	1059	10287	1179	995	56	12517
2	3496	4650	970	596	7760	952	928	72	9712
3- 4	5323	7467	1312	794	11479	1761	1567	89	14896
5- 8	7165	10185	1587	806	15076	2607	1896	164	19743
9-16	9228	13771	2037	936	19506	3680	2565	221	25972
17-32	12669	17696	1634	881	24028	5081	3318	453	32880
33-64	17371	20481	1241	822	26548	7381	5354	632	39915
65-	161665	90500	2114	2392	140467	74935	35342	5927	256671
計	221399	170159	12462	8286	255151	97576	51965	7614	412306

	人	地	総計
1	4280	1269	18066
2	2840	610	13162
3- 4	3492	785	19173
5- 8	3209	829	23781
9-16	2272	864	29108
17-32	1347	1024	35251
33-64	714	717	41346
65-	69	2133	258873
計	18223	8231	438760

つぎに、上位何語までで、どのくらいの使用度数をカバーするかをしめす。

順位	和	漢	外	混	人	地	のべ計	カバー率
100	65	33	1	0	0	1	143696	0.328
200	130	63	2	2	0	3	176864	0.403
300	182	108	3	4	0	3	197671	0.451
400	235	150	3	7	0	5	212710	0.485
500	291	191	5	8	0	5	224828	0.512
600	341	236	8	9	0	6	234975	0.536
700	385	286	11	12	0	6	243725	0.555
800	425	342	14	13	0	6	251431	0.573
900	471	389	15	15	1	9	258288	0.589
1000	515	439	16	16	4	10	264446	0.603
....								
2000	905	954	50	38	24	29	305675	0.697
3000	1268	1456	98	68	53	57	329090	0.750
4000	1604	1942	157	95	114	88	345045	0.786
5000	1945	2375	230	126	198	126	356816	0.813
6000	2227	2858	309	155	300	151	366010	0.834
7000	2522	3292	371	199	431	185	373449	0.851
8000	2828	3747	430	221	552	222	379666	0.865
9000	3152	4130	508	260	692	258	384924	0.877
10000	3426	4559	570	289	873	283	389611	0.888
....								
15000	4827	6521	889	489	1780	494	406482	0.926

20000	6153	8202	1330	734	2857	724	416832	0.950
25000	7546	9822	1618	983	3982	1049	423763	0.966
30000	9034	11106	2011	1295	5139	1415	428763	0.977
35000	9859	13191	2455	1575	6191	1729	433763	0.989
39997	11152	14369	2988	1868	7539	2081	438760	1.000

4. 表記のゆれの統計

言語処理データ集7には、単語ごとに、ちがった表記の度数をしるした。ただし、この<表記のゆれ>には、「やはり～やっぱり」「イヤリング～イヤリング」のような語形自体のゆれもふくむ。カタカナとひらがなの区別、おくりがなの差は、別表記としてかぞえたが、歴史的かなづかいは現代かなづかいに、異体の漢字は通用の字体の漢字に合併した。

この表については、どのような<ゆれ>がみとめられるか、という分析をすべきであるが、それは大きなしごとになるので、ここでは、どのような単語にゆれがみられるか、ということのべておく。まず、度数順にみると、当然度数1のものにはゆれがなく、度数のたかいものほど、これがおおい。以下、「ゆれ」としたのは表記のゆれのみられる語数である。

度数	2	3-4	5-8	9-16	17-32	33-64	65-
ゆれ	1035	1348	1225	871	645	434	580
全体	6581	5663	3866	2462	1536	914	909
比率	.157	.238	.317	.354	.420	.475	.638

つぎに、語種・品詞別にみると、

	和	漢	外	混	名	動	形	他	人	地
ゆれ	4210	1071	322	236	3149	1657	923	110	206	93
全体	11152	14369	2988	1868	23862	3457	2839	219	7539	2081
比率	.378	.075	.108	.126	.132	.479	.325	.502	.027	.045

で、ゆれがおおいは和語・動詞、すくないのは漢語・固有名詞である。

表記形のおおい例（表記数14以上）を、つぎにあげる。これは、表記表の見本でもある。各レコードは、つぎの4つのフィールドからなる。

見出し、語種・品詞、合計度数、表記とその度数

見出し・語種・表記は全角（2バイト）で、それ以外は半角（1バイト）で記入してある。各フィールドは半角のコンマで、各表記のあいだは半角の空白で、くぎったが、表記とその度数のあいだは、くぎらずにつづけた。（ただし、表記の種類が1つのばあいは、表記別の度数を省略した。）

表記数23) ナニ,和1,1263, あん(も)1 なーに2 なあに4 なあん3 なに138 なにッ1

なん229 なんに8 なアーン1 なアに2 なアン2 なん1 ナーニ1 ナン1 ナンに2
 何831 何(ナン)に2 何あん1 何っ1 何に7 何ん21 何ン3 甚(なに)1
 表記数15) ヨイ,和3,1601, い>49 いい515 え>2 ええ17 よい910 よう(こそ)10
 ようく1 エー1 ヨイ2 佳い1 宜い1 好い7 善い3 能い4 良い78
 表記数14) エ, エエ,和4,77, \ / 1 え18 え>10 えーつと1 えいッ1 ええ34 ええい1
 ええっ2 ええッ2 えっ1 えッ3 エ>1 エイッ1 エエイッ1
 表記数13) イウ,和2,6147, いう5086 ちゅう3 って40 ってえ8 とう1 チュ1 テ2
 謂う1 云う455 言う547 言ウ1 雖も1 雖モ1

表記数13) ソレ,和1, 2172, そいで2 そりゃ18 そりゃあ6 そりゃア2 それ2122
 それあ1 それっ4 それやあ1 それア1 そんなら10 そんなら1 ソレ3 夫1

上位には、和語がならぶ。漢語で変種がいちばんおおいのは、

表記数10) ゴセン,漢1,150, 5{・・・}17 5{億}4 5{千}2 50{百}17 5000 7 5千1
 五〇〇〇 2 五{・・・}9 五{千}20 五千71

であり、原文には単に「5、五」などとあるのを、くらいどりをはっきりさせるために読みそえ、これを別表記とみなしたものである。数詞は度数がたかいこともあって、漢語の多表記語には、この種のものおおい。

外来語のトップにくるのは、

表記数5) ショオウインドオ,外英1,5, ショー・ウィンドー1 ショーウィンドー1
 ショーウインドー1 ショウ・ウィンドー1 ショウ・ウィンドウ1

であり、地名では

表記数5) イタリア,地,42, イタリー4 イタリア32 イタリイ1 イタリヤ3 伊太利2

表記数5) ソビエト,地,19, ソビエト1 ソヴィエト2 ソヴェート2 ソヴェト4
 ソヴェト10

が、人名では

表記数4) ジェエムズ,名,11, ジェームス6 ジェームズ2 ジェイムズ2 ジェイムズ1

が最高だった。

つぎに、ある語の表記数と使用度数の関係を表にしてしめす。

度数 表記	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16以上	計
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	3
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	6
10	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	4	5
9	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	5	6
8	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1	19	21
7	-	-	-	-	-	-	1	-	-	-	-	1	-	1	1	28	32
6	-	-	-	-	-	-	-	1	2	-	2	1	-	2	-	51	59
5	-	-	-	-	8	6	9	12	6	3	7	5	1	3	1	119	180
4	-	-	-	25	30	24	20	19	15	14	12	11	10	10	3	247	440
3	-	-	147	135	94	110	66	60	45	42	46	31	23	30	13	504	1346
2	-	1035	654	387	303	188	159	115	84	82	80	61	46	51	46	744	4035
1	18066	5546	2678	1637	994	713	523	411	296	303	206	202	178	147	134	1825	33859
計	18066	6581	3479	2184	1429	1041	778	618	448	444	354	312	259	245	199	3560	39997

一方、使用度数がたかいのに表記にゆれの少ない語には、どのようなものがあるか。和語では、この調査で自立語あつかいにした、

れる・られる (3521), たい (662), せる・させる (487)

など、漢語では

年 (1337), 円 (1031), 月 (944), 第 (740), 者 (689), 人 (548)

などの接頭接尾的要素が上位にくる。これらをのぞいた一般的な用語では、

[和語] 考える (616), そして (438), 男 (417), 身頃 (349), 日 (339), そんな (339)

[漢語] 問題 (468), 映画 (407), 生活 (389), 気 (351), 必要 (267), 会社 (267)

[外来語] スカート (123), デザイン (81), ダーツ (74), バス (58), ポケット (57)

[人名] 山本 (69), 田中 (64), 中村 (61), 鈴木 (59), 林 (53), 木村 (53), 吉田 (47)

[地名] 大阪 (149), ソ連 (142), 京都 (68), 中国 (64), 北海道 (58), 米国 (56)

などがある。

さて、ことなり語数39,997のなかには、度数1のものが18,066ふくまれているから、これをいれると、ゆれのある語数6,138は、全体の15.3%にすぎない。もっとも、ゆれのある語のおおくは使用度数がたかいから、のべ語数でいうと、438,760語中の264,730語、つまり約60.3%になる。文章中の6割の語は、なにか、ほかの語形をもっている可能性があるようにみえる。しかし、この

推定は現実的ではない。ゆれている、といっても、おおくの語は、一部の表記に集中しているのである。度数のたかいもので、特定の形式に99%以上が集中している例をあげると、

- 99.7%) コウ,和3,390,こう389 斯う1
- 99.4%) ニッポン,地,890,にほん2 ニッポン3 日本885
- 99.3%) チチ,和1,165,チチ1 父164
- 99.2%) アル,和2,3160,ある3136 アル6 在る12 有る6
- 99.2%) センセイ,漢1,261,せんせい1 センセ1 先生259

などがある。これらも、ここではくゆれのある>ものに分類しているのである。いま、ひとつの形式しかないものをく固定>、特定の形式が 9割以上をしめているものをく独占>、それ以外をくゆれ>とよぶとすれば、く独占>は使用度数のたかいものにおおくみられる。

度数	2	3-4	5-8	9-16	17-32	33-64	65-	合計
ゆれ	1035	1348	1225	694	462	299	339	5402
独占	-	-	-	177	183	135	241	736
固定	5546	4315	2641	1591	891	480	329	

このく独占>をく固定>にふくめると、くゆれ>グループののべ語数は、152,122語 (34.7%) におちる。つまり、約3割の語の表記がゆれる、ということで、このほうが実際の感じにちかいかであろう。

参考文献

- 国立国語研究所 (1962) 『現代雑誌九十種の用語用字 第一分冊 総記および語彙表』(国立国語研究所報告21)
- 国立国語研究所 (1963) 『現代雑誌九十種の用語用字 第二分冊 漢字表』(国立国語研究所報告22)
- 国立国語研究所 (1964) 『現代雑誌九十種の用語用字 第三分冊 分析』(国立国語研究所報告25)
- 国立国語研究所 (1987) 『現代雑誌九十種の用語用字 五十音順語彙表・採集カード』(国立国語研究所言語処理データ集3, マイクロフィッシュ)

(原稿受理日：1997年2月14日)

宮島 達夫 (みやじま たつお)

京都橘女子大学文学部国文学科 607 京都市山科区大宅山田町34
 SGQ02466@niftyserve.or.jp

Statistical analysis of written word-forms in ninety magazines

MIYAZIMA Tatu
Kyoto Tachibana Women's University

Key words

National Language Research Institute, magazine vocabulary, vocabulary
research, variation in written forms

The National Language Research Institute carried out a large-scale statistical study of the language of ninety magazines of 1956. The results were published in a series of reports. The data will soon be made public in a floppy disk including frequency of each written word-form. In this paper some statistical tables were revised based on the new data and some tables were added concerning variation in word-forms. It is pointed out that native Japanese words and verbs have more varieties in form while Chinese loans and proper nouns have fewer varieties.